



OIPE

## RAW SEQUENCE LISTING

DATE: 02/11/2002

PATENT APPLICATION: US/09/866,379

TIME: 09:07:06

Input Set : A:\DIVER1370-7.ST25.txt

Output Set: N:\CRF3\02112002\I866379.raw

P.S.  
**ENTERED**

```

2 <110> APPLICANT: DIVERSA CORPORATION
3     SHORT, Jay
4     KRETZ, Keith
5     GRAY, Kevin
6     BARTON, Nelson
7     GARRETT, James
8     O'DONOGHUE, Eileen
10 <120> TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
12 <130> FILE REFERENCE: DIVER1370-7
14 <140> CURRENT APPLICATION NUMBER: US 09/866,379
15 <141> CURRENT FILING DATE: 2001-05-24
17 <150> PRIOR APPLICATION NUMBER: US 09/580,515
18 <151> PRIOR FILING DATE: 2000-05-25
20 <150> PRIOR APPLICATION NUMBER: US 09/318,528
21 <151> PRIOR FILING DATE: 1999-05-25
23 <150> PRIOR APPLICATION NUMBER: US 09/291,931
24 <151> PRIOR FILING DATE: 1999-04-13
26 <150> PRIOR APPLICATION NUMBER: US 09/259,214
27 <151> PRIOR FILING DATE: 1999-03-01
29 <150> PRIOR APPLICATION NUMBER: US 08/910,798
30 <151> PRIOR FILING DATE: 1997-08-13
32 <160> NUMBER OF SEQ ID NOS: 10
34 <170> SOFTWARE: PatentIn version 3.1
36 <210> SEQ ID NO: 1
37 <211> LENGTH: 1323
38 <212> TYPE: DNA
39 <213> ORGANISM: Escherichia coli
41 <220> FEATURE:
42 <221> NAME/KEY: misc_feature
43 <222> LOCATION: (1)..(1323)
44 <223> OTHER INFORMATION: n is any nucleotide
46 <220> FEATURE:
47 <221> NAME/KEY: CDS
48 <222> LOCATION: (1)..(1323)
49 <223> OTHER INFORMATION:
51 <400> SEQUENCE: 1
52 atg aaa gcg atc tta atc cca ttt tta tct ctt ctg att ccg tta acc      48
53 Met Lys Ala Ile Leu Ile Pro Phe Leu Ser Leu Leu Ile Pro Leu Thr
54 1          5          10          15
56 ccg caa tct gca ttc gct cag agt gag ccg gag ctg aag ctg gaa agt      96
57 Pro Gln Ser Ala Phe Ala Gln Ser Glu Pro Glu Leu Lys Leu Glu Ser
58          20          25          30
60 gtg gtg att gtc agt cgt cat ggt gtg cgt gct cca acc aag gcc acg      144

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/866,379

DATE: 02/11/2002

TIME: 09:07:06

Input Set : A:\DIVER1370-7.ST25.txt

Output Set: N:\CRF3\02112002\I866379.raw

```

61 Val Val Ile Val Ser Arg His Gly Val Arg Ala Pro Thr Lys Ala Thr
62          35          40          45
64 caa ctg atg cag gat gtc acc cca gac gca tgg cca acc tgg ccg gta      192
65 Gln Leu Met Gln Asp Val Thr Pro Asp Ala Trp Pro Thr Trp Pro Val
66          50          55          60
W--> 68 aaa ctg ggt tgg ctg aca ccg cgn ggt ggt gag cta atc gcc tat ctc      240
69 Lys Leu Gly Trp Leu Thr Pro Arg Gly Gly Glu Leu Ile Ala Tyr Leu
70 65          70          75          80
72 gga cat tac caa cgc cag cgt ctg gta gcc gac gga ttg ctg gcg aaa      288
73 Gly His Tyr Gln Arg Gln Arg Leu Val Ala Asp Gly Leu Leu Ala Lys
74          85          90          95
76 aag ggc tgc ccg cag tct ggt cag gtc gcg att att gct gat gtc gac      336
77 Lys Gly Cys Pro Gln Ser Gly Gln Val Ala Ile Ile Ala Asp Val Asp
78          100          105          110
80 gag cgt acc cgt aaa aca ggc gaa gcc ttc gcc gcc ggg ctg gca cct      384
81 Glu Arg Thr Arg Lys Thr Gly Glu Ala Phe Ala Ala Gly Leu Ala Pro
82          115          120          125
84 gac tgt gca ata acc gta cat acc cag gca gat acg tcc agt ccc gat      432
85 Asp Cys Ala Ile Thr Val His Thr Gln Ala Asp Thr Ser Ser Pro Asp
86          130          135          140
88 ccg tta ttt aat cct cta aaa act ggc gtt tgc caa ctg gat aac gcg      480
89 Pro Leu Phe Asn Pro Leu Lys Thr Gly Val Cys Gln Leu Asp Asn Ala
90 145          150          155          160
92 aac gtg act gac gcg atc ctc agc agg gca gga ggg tca att gct gac      528
93 Asn Val Thr Asp Ala Ile Leu Ser Arg Ala Gly Gly Ser Ile Ala Asp
94          165          170          175
96 ttt acc ggg cat ccg caa acg gcg ttt cgc gaa ctg gaa ccg gtg ctt      576
97 Phe Thr Gly His Arg Gln Thr Ala Phe Arg Glu Leu Glu Arg Val Leu
98          180          185          190
100 aat ttt ccg caa tca aac ttg tgc ctt aaa cgt gag aaa cag gac gaa      624
101 Asn Phe Pro Gln Ser Asn Leu Cys Leu Lys Arg Glu Lys Gln Asp Glu
102          195          200          205
104 agc tgt tca tta acg cag gca tta cca tcg gaa ctc aag gtg agc gcc      672
105 Ser Cys Ser Leu Thr Gln Ala Leu Pro Ser Glu Leu Lys Val Ser Ala
106          210          215          220
108 gac aat gtc tca tta acc ggt gcg gta agc ctc gca tca atg ctg acg      720
109 Asp Asn Val Ser Leu Thr Gly Ala Val Ser Leu Ala Ser Met Leu Thr
110 225          230          235          240
112 gag ata ttt ctc ctg caa caa gca cag gga atg ccg gag ccg ggg tgg      768
113 Glu Ile Phe Leu Leu Gln Gln Ala Gln Gly Met Pro Glu Pro Gly Trp
114          245          250          255
116 gga agg atc acc gat tca cac cag tgg aac acc ttg cta agt ttg cat      816
117 Gly Arg Ile Thr Asp Ser His Gln Trp Asn Thr Leu Leu Ser Leu His
118          260          265          270
120 aac gcg caa ttt tat ttg cta caa cgc acg cca gag gtt gcc cgc agc      864
121 Asn Ala Gln Phe Tyr Leu Leu Gln Arg Thr Pro Glu Val Ala Arg Ser
122          275          280          285
124 cgc gcc acc ccg tta ttg gat ttg atc atg gca gcg ttg acg ccc cat      912
125 Arg Ala Thr Pro Leu Leu Asp Leu Ile Met Ala Ala Leu Thr Pro His

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/866,379

DATE: 02/11/2002

TIME: 09:07:07

Input Set : A:\DIVER1370-7.ST25.txt

Output Set: N:\CRF3\02112002\I866379.raw

```

126      290      295      300
128 cca ccg caa aaa cag gcg tat ggt gtg aca tta ccc act tca gta ctg      960
129 Pro Pro Gln Lys Gln Ala Tyr Gly Val Thr Leu Pro Thr Ser Val Leu
130 305      310      315      320
132 ttt att gcc gga cac gat act aat ctg gca aat ctc ggc ggc gca ctg      1008
133 Phe Ile Ala Gly His Asp Thr Asn Leu Ala Asn Leu Gly Gly Ala Leu
134      325      330      335
136 gag ctc aac tgg acg ctt ccc ggt cag ccg gat aac acg ccg cca ggt      1056
137 Glu Leu Asn Trp Thr Leu Pro Gly Gln Pro Asp Asn Thr Pro Pro Gly
138      340      345      350
140 ggt gaa ctg gtg ttt gaa cgc tgg cgt cgg cta agc gat aac agc cag      1104
141 Gly Glu Leu Val Phe Glu Arg Trp Arg Arg Leu Ser Asp Asn Ser Gln
142      355      360      365
144 tgg att cag gtt tcg ctg gtc ttc cag act tta cag cag atg cgt gat      1152
145 Trp Ile Gln Val Ser Leu Val Phe Gln Thr Leu Gln Gln Met Arg Asp
146      370      375      380
148 aaa acg ccg ctg tca tta aat acg ccg ccc gga gag gtg aaa ctg acc      1200
149 Lys Thr Pro Leu Ser Leu Asn Thr Pro Pro Gly Glu Val Lys Leu Thr
150 385      390      395      400
152 ctg gca gga tgt gaa gag cga aat gcg cag ggc atg tgt tcg ttg gca      1248
153 Leu Ala Gly Cys Glu Glu Arg Asn Ala Gln Gly Met Cys Ser Leu Ala
154      405      410      415
156 ggt ttt acg caa atc gtg aat gaa gca cgc ata ccg gcg tgc agt ttg      1296
157 Gly Phe Thr Gln Ile Val Asn Glu Ala Arg Ile Pro Ala Cys Ser Leu
158      420      425      430
160 aga tct cat cac cat cac cat cac taa      1323
161 Arg Ser His His His His His His
162      435      440
165 <210> SEQ ID NO: 2
166 <211> LENGTH: 440
167 <212> TYPE: PRT
168 <213> ORGANISM: Escherichia coli
170 <400> SEQUENCE: 2
172 Met Lys Ala Ile Leu Ile Pro Phe Leu Ser Leu Leu Ile Pro Leu Thr
173 1      5      10      15
176 Pro Gln Ser Ala Phe Ala Gln Ser Glu Pro Glu Leu Lys Leu Glu Ser
177      20      25      30
180 Val Val Ile Val Ser Arg His Gly Val Arg Ala Pro Thr Lys Ala Thr
181      35      40      45
184 Gln Leu Met Gln Asp Val Thr Pro Asp Ala Trp Pro Thr Trp Pro Val
185      50      55      60
188 Lys Leu Gly Trp Leu Thr Pro Arg Gly Gly Glu Leu Ile Ala Tyr Leu
189 65      70      75      80
192 Gly His Tyr Gln Arg Gln Arg Leu Val Ala Asp Gly Leu Leu Ala Lys
193      85      90      95
196 Lys Gly Cys Pro Gln Ser Gly Gln Val Ala Ile Ile Ala Asp Val Asp
197      100      105      110
200 Glu Arg Thr Arg Lys Thr Gly Glu Ala Phe Ala Ala Gly Leu Ala Pro
201      115      120      125

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/866,379

DATE: 02/11/2002

TIME: 09:07:07

Input Set : A:\DIVER1370-7.ST25.txt

Output Set: N:\CRF3\02112002\I866379.raw

```

204 Asp Cys Ala Ile Thr Val His Thr Gln Ala Asp Thr Ser Ser Pro Asp
205      130      135      140
208 Pro Leu Phe Asn Pro Leu Lys Thr Gly Val Cys Gln Leu Asp Asn Ala
209 145      150      155      160
212 Asn Val Thr Asp Ala Ile Leu Ser Arg Ala Gly Gly Ser Ile Ala Asp
213      165      170      175
216 Phe Thr Gly His Arg Gln Thr Ala Phe Arg Glu Leu Glu Arg Val Leu
217      180      185      190
220 Asn Phe Pro Gln Ser Asn Leu Cys Leu Lys Arg Glu Lys Gln Asp Glu
221      195      200      205
224 Ser Cys Ser Leu Thr Gln Ala Leu Pro Ser Glu Leu Lys Val Ser Ala
225      210      215      220
228 Asp Asn Val Ser Leu Thr Gly Ala Val Ser Leu Ala Ser Met Leu Thr
229 225      230      235      240
232 Glu Ile Phe Leu Leu Gln Gln Ala Gln Gly Met Pro Glu Pro Gly Trp
233      245      250      255
236 Gly Arg Ile Thr Asp Ser His Gln Trp Asn Thr Leu Leu Ser Leu His
237      260      265      270
240 Asn Ala Gln Phe Tyr Leu Leu Gln Arg Thr Pro Glu Val Ala Arg Ser
241      275      280      285
244 Arg Ala Thr Pro Leu Leu Asp Leu Ile Met Ala Ala Leu Thr Pro His
245      290      295      300
248 Pro Pro Gln Lys Gln Ala Tyr Gly Val Thr Leu Pro Thr Ser Val Leu
249 305      310      315      320
252 Phe Ile Ala Gly His Asp Thr Asn Leu Ala Asn Leu Gly Gly Ala Leu
253      325      330      335
256 Glu Leu Asn Trp Thr Leu Pro Gly Gln Pro Asp Asn Thr Pro Pro Gly
257      340      345      350
260 Gly Glu Leu Val Phe Glu Arg Trp Arg Arg Leu Ser Asp Asn Ser Gln
261      355      360      365
264 Trp Ile Gln Val Ser Leu Val Phe Gln Thr Leu Gln Gln Met Arg Asp
265      370      375      380
268 Lys Thr Pro Leu Ser Leu Asn Thr Pro Pro Gly Glu Val Lys Leu Thr
269 385      390      395      400
272 Leu Ala Gly Cys Glu Glu Arg Asn Ala Gln Gly Met Cys Ser Leu Ala
273      405      410      415
276 Gly Phe Thr Gln Ile Val Asn Glu Ala Arg Ile Pro Ala Cys Ser Leu
277      420      425      430
280 Arg Ser His His His His His His
281      435      440
284 <210> SEQ ID NO: 3
285 <211> LENGTH: 49
286 <212> TYPE: DNA
287 <213> ORGANISM: Artificial Sequence
289 <220> FEATURE:
290 <223> OTHER INFORMATION: Primer for PCR
292 <400> SEQUENCE: 3
293 gtttctgaat tcaaggagga atttaaataa aagcgatctt aatcccat
296 <210> SEQ ID NO: 4

```

49

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/866,379

DATE: 02/11/2002

TIME: 09:07:07

Input Set : A:\DIVER1370-7.ST25.txt

Output Set: N:\CRF3\02112002\I866379.raw

```

297 <211> LENGTH: 33
298 <212> TYPE: DNA
299 <213> ORGANISM: Artificial Sequence
301 <220> FEATURE:
302 <223> OTHER INFORMATION: Primer for PCR
304 <400> SEQUENCE: 4
305 gtttctggat ccttacaac tgcacgccg tat
308 <210> SEQ ID NO: 5
309 <211> LENGTH: 1901
310 <212> TYPE: DNA
311 <213> ORGANISM: Escherichia coli
313 <220> FEATURE:
314 <221> NAME/KEY: misc_feature
315 <222> LOCATION: (1)..(1901)
316 <223> OTHER INFORMATION: n is any nucleotide
318 <400> SEQUENCE: 5
319 taaggagcag aaacaatgtg gtatttactt tggttcgctg gcattttgtt gatgtgttcg 60
321 ctctccaccc ttgtgttggg atggctggac ccgcgtctga aaagttaacg aacgtaggcc 120
323 tgatgcggcg cattagcatt gcacaggca atcaataatg tcagatatga aaagcggaaa 180
325 catatcgatg aaagcgatct taatcccat tttatctctt ctgattccgt taaccccgca 240
327 atctgcattc gctcagagtg agccggagct gaagctggaa agtgtggtga ttgtcagtcg 300
329 tcatgggtgtg cgtgctccaa ccaaggccac gcaactgatg caggatgtca cccagacgc 360
W--> 331 atggccaacc tggccggtaa aactgggtg actgacaccg cngngtggtg agctaatacgc 420
333 ctatctcgga cattaccaac gccagcgtct ggtagccgac ggattgctgg cgaaaaaggg 480
335 ctgcccgcag tctggtcagg tcgcgattat tgctgatgtc gacgagcgta cccgtaaac 540
337 aggcgaagcc ttgcgcgcgc ggctggcacc tgaactgtga ataaccgtac ataccaggc 600
339 agatacgtcc agtcccgcgc cgttatttaa tccctaaaaa actggcggtt gccaaactgga 660
341 taacgcgaac gtgactgacg cgatcctcag cagggcagga gggtaaatg ctgactttac 720
343 cgggcagtcg caaacggcgt ttgcgcaact ggaacgggtg cttaatcttc cgcaatcaaa 780
345 cttgtgcctt aaacgtgaga aacaggacga aagctgttca ttaacgcagg cattaccatc 840
347 ggaactcaag gtgagcgcgc acaatgtctc attaacgggt gcgtaagcc tcgcatcaat 900
349 gctgacggag atatttctcc tgcaacaagc acaggaatg ccggagccgg ggtggggaag 960
351 gateaccgat tcacaccagt ggaacacctt gctaagtttg cataacgcgc aattttatct 1020
353 gctacaacgc acgccagagg ttgcccgcag ccgcgccacc ccgttattag atttgatcaa 1080
355 gacagcgttg acgccccatc caccgcaaaa acaggcgtat ggtgtgacat taccacttc 1140
357 agtgcgtgtt atcgccggac acgatactaa tctggcaaat ctgcgcggcg cactggagct 1200
359 caactggacg cttcccgttc agccggataa cagccgcga ggtggtgaac tgggtgttga 1260
361 acgctggcgt cggctaagcg ataacagcca gtggattcag gtttcgctgg tcttcagac 1320
363 ttacagcag atgcgtgata aaacgcgcgt gtcattaaat acgcgcgccg gagaggtgaa 1380
365 actgaccctg gcaggatgtg aagagcgaaa tgcgcagggc atgtgttcgt tggcagggtt 1440
367 tacgcaaate gtgaatgaag cacgcatacc ggcgtgcagt ttgtaatgca taaaaaagag 1500
369 cattcagtta cctgaatgct ctgaggctga tgacaaacga agaactgtct aatgcgtaga 1560
371 ccggaagagg cgttcacgcc gcacccggcc actttcagtt ttctcttttc tcggagtaac 1620
373 tataaccgta atagtatat ccgtaactgt aagcgggtgt ggcgcgttta atcacaccat 1680
375 tgaggatagc gcctttaata ttgacgcctg cctgttcag acgctgcatt gacaaactca 1740
377 cctctttggc ggtgttcaag ccaaacgcgc caaccagcag gctggtgcca acagaacgcc 1800
379 ccacgaccgc ggcacactc accgccagca tcggcggcgt atcgacaate accagatcgt 1860
381 aatggtcggt cggccattcc agtaattgac gcacccgacg g
384 <210> SEQ ID NO: 6

```

—)

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/866,379

DATE: 02/11/2002

TIME: 09:07:08

Input Set : A:\DIVER1370-7.ST25.txt

Output Set: N:\CRF3\02112002\I866379.raw

L:68 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:331 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:407 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
L:483 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:674 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9